# Web Supplement to *Effects of adjusting for instrumental variables on bias and precision of effect estimates*

October 3, 2011

## Web Appendix 1: Code

All data generation, analysis, and plotting was performed in R. In this section, we provide the code used for simulation so that others may reproduce our results. The function `addiSims`, along with the accompanying functions `rd.crude` and `rd.cond`, simulates and analyzes data for one set of simulation parameters in the additive framework. The function `multiSims`, along with the accompanying functions `rr.crude` and `rr.cond`, simulates and analyzes data for one set of simulation parameters in the multiplicative framework.

```
rd.cond <- function(y, x, c) {
  cases1 <- rowsum(y*x, c)
  cases0 <- rowsum(y*(1-x), c)
  n1 <- rowsum(x, c)
  n0 <- rowsum(1-x, c)
  n <- c(sum(1-c), sum(c))
  sum((cases1*n0 - cases0*n1)/n) / sum(n1*n0/n)
}

rd.crude <- function(y, x) {
  n1 <- sum(x)
  n0 <- sum(1-x)
  p1 <- sum(y*x)/n1
  p0 <- sum(y*(1-x))/n0
  p1 - p0
}
```

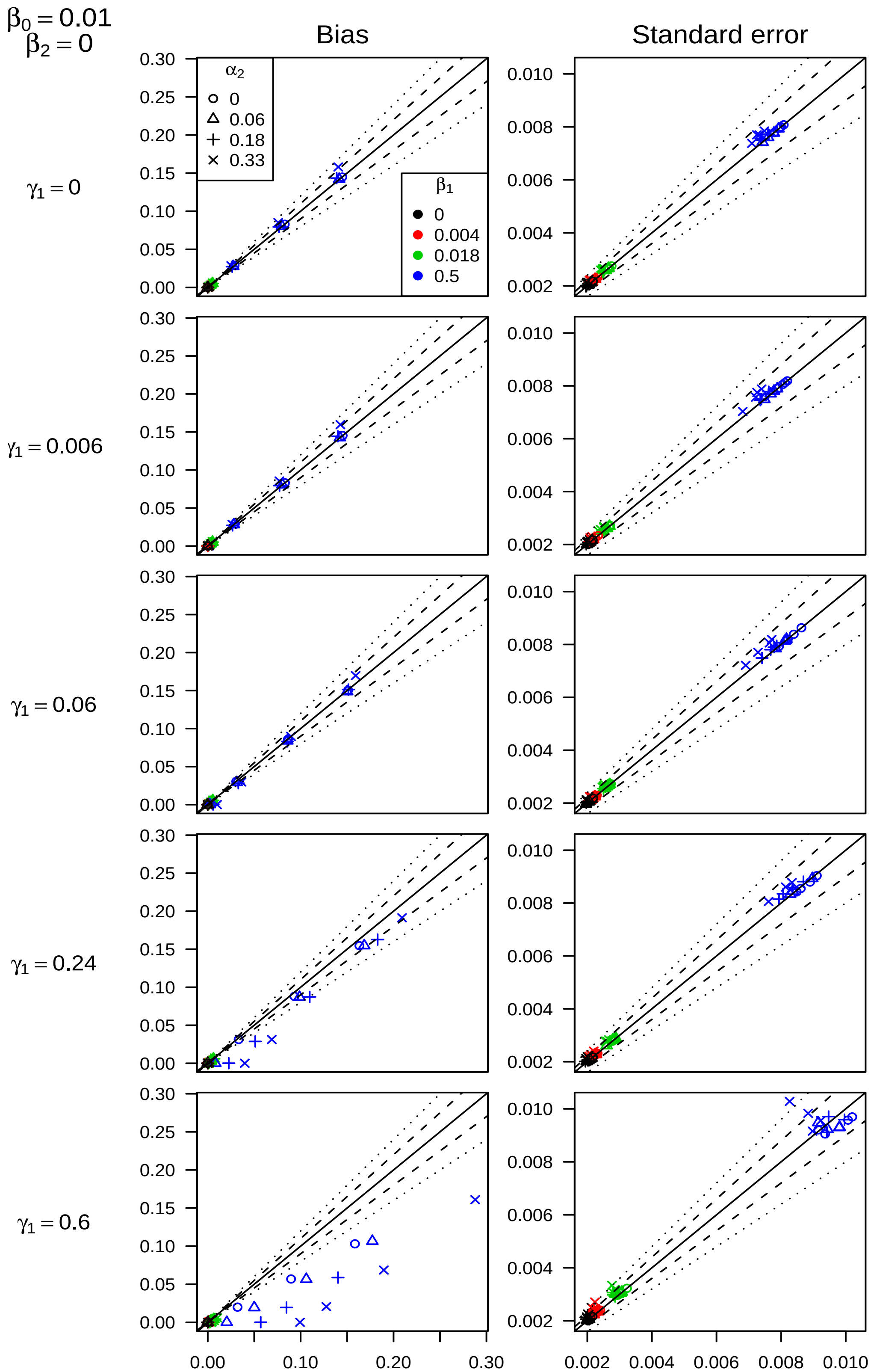
```
addiSims <- function(simpars, nsamp=10000, nsim=2500) {

  results <- matrix(0, nsim, 3)

  dat <- matrix(NA, nsim, 2^3)
  colnames(dat) <- c("z0x0y0", "z0x0y1", "z0x1y0", "z0x1y1",
                     "z1x0y0", "z1x0y1", "z1x1y0", "z1x1y1")

  # simpars should be an R data frame or list with named elements gamma0-beta2
  g0 <- simpars$gamma0
  g1 <- simpars$gamma1
  a0 <- simpars$alpha0
  a1 <- simpars$alpha1
  a2 <- simpars$alpha2
  b0 <- simpars$beta0
  b1 <- simpars$beta1
  b2 <- simpars$beta2

  for(s in 1:nsim){

                                      # make the data
    z <- rbinom(nsamp, 1, .5)
    u <- rbinom(nsamp, 1, g0 + g1*z)
    x <- rbinom(nsamp, 1, a0 + a1*u + a2*z)
    y <- rbinom(nsamp, 1, b0 + b1*u + b2*x)
    dat[s,] <- as.vector(table(y,x,z))

                                      # estimates
    results[s,] <- round(c(rd.crude(y, x), # unadjusted association
                           rd.cond(y, x, u), # adjusting for u
                           rd.cond(y, x, z), # adjusting for z
                          ), 6)
  }
  colnames(results) <- c("crude", "truth", "condZ")

  results <- cbind(results, dat)
  results
}

rr.cond <- function(y, x, c) {
  cases1 <- rowsum(y*x, c)
  cases0 <- rowsum(y*(1-x), c)
  n1 <- rowsum(x, c)
  n0 <- rowsum(1-x, c)
```

```
  n <- c(sum(1-c), sum(c))
  sum(cases1*n0/n)/sum(cases0*n1/n)
}

rr.crude <- function(y, x) {
  n1 <- sum(x)
  n0 <- sum(1-x)
  p1 <- sum(y*x)/n1
  p0 <- sum(y*(1-x))/n0
  p1/p0
}

multiSims <- function(simpars, nsamp=10000, nsim=2500) {

  results <- matrix(0, nsim, 3)

  dat <- matrix(NA, nsim, 2^3)
  colnames(dat) <- c("z0x0y0", "z0x0y1", "z0x1y0", "z0x1y1",
                     "z1x0y0", "z1x0y1", "z1x1y0", "z1x1y1")

  # simpars should be an R data frame or list with named elements gamma0-beta2
  g0 <- simpars$gamma0
  g1 <- simpars$gamma1
  a0 <- simpars$alpha0
  a1 <- simpars$alpha1
  a2 <- simpars$alpha2
  b0 <- simpars$beta0
  b1 <- simpars$beta1
  b2 <- simpars$beta2

  for(s in 1:nsim){

                                     # make the data
    z <- rbinom(nsamp, 1, .5)
    u <- rbinom(nsamp, 1, g0 * g1^z)
    x <- rbinom(nsamp, 1, a0 * a1^u * a2^z)
    y <- rbinom(nsamp, 1, b0 * b1^u * b2^x)
    dat[s,] <- as.vector(table(y,x,z))

                                     # estimates
    results[s,] <- round(c(rr.crude(y, x), # unadjusted association
                           rr.cond(y, x, u), # adjusted for u
                           rr.cond(y, x, z), # adjusted for z
                           ), 6)
  }
```

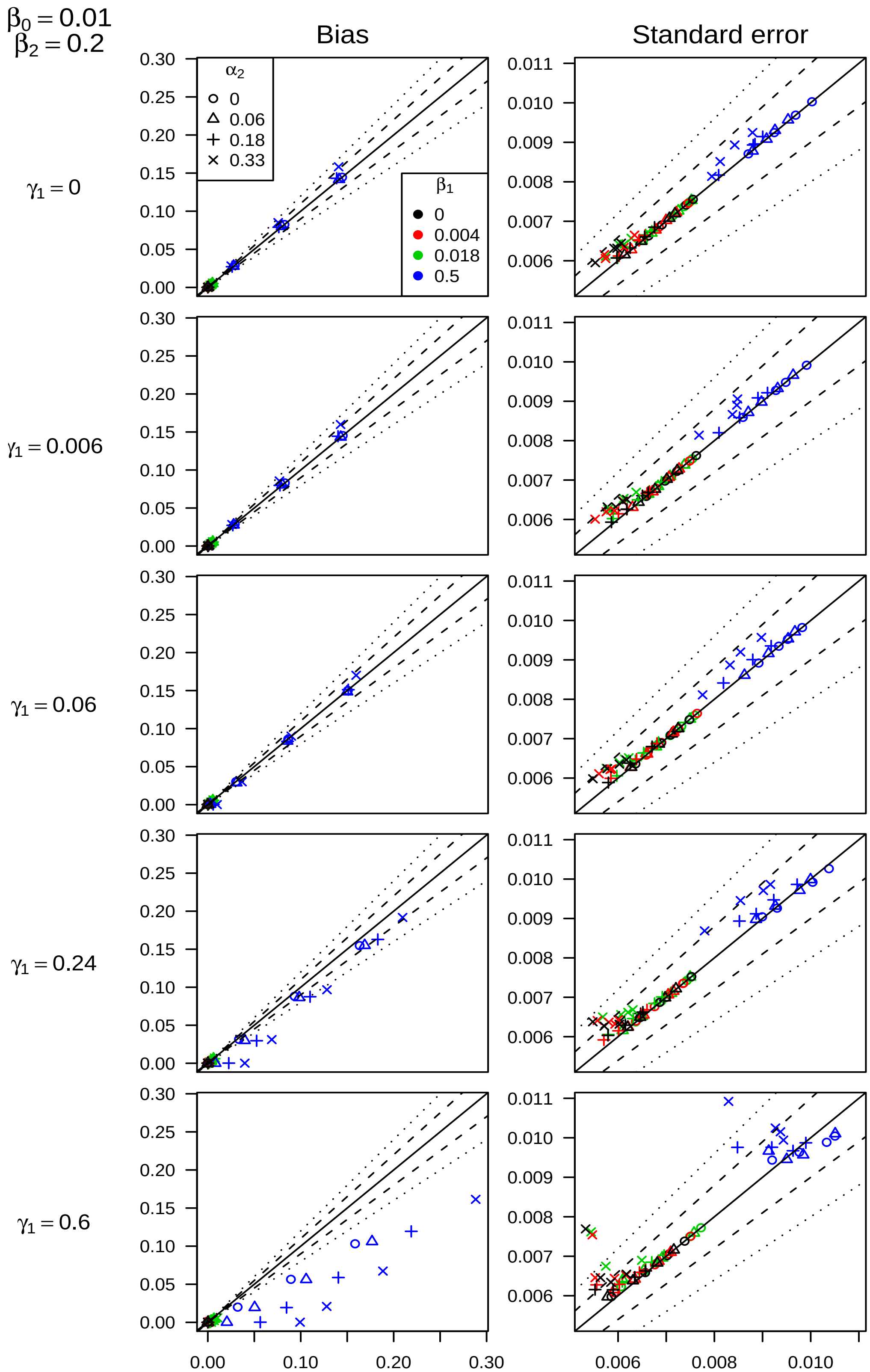
```
  colnames(results) <- c("crude", "truth", "condZ")

  results <- cbind(results, dat)
  results
}
```

## Web Appendix 2: Simulation results

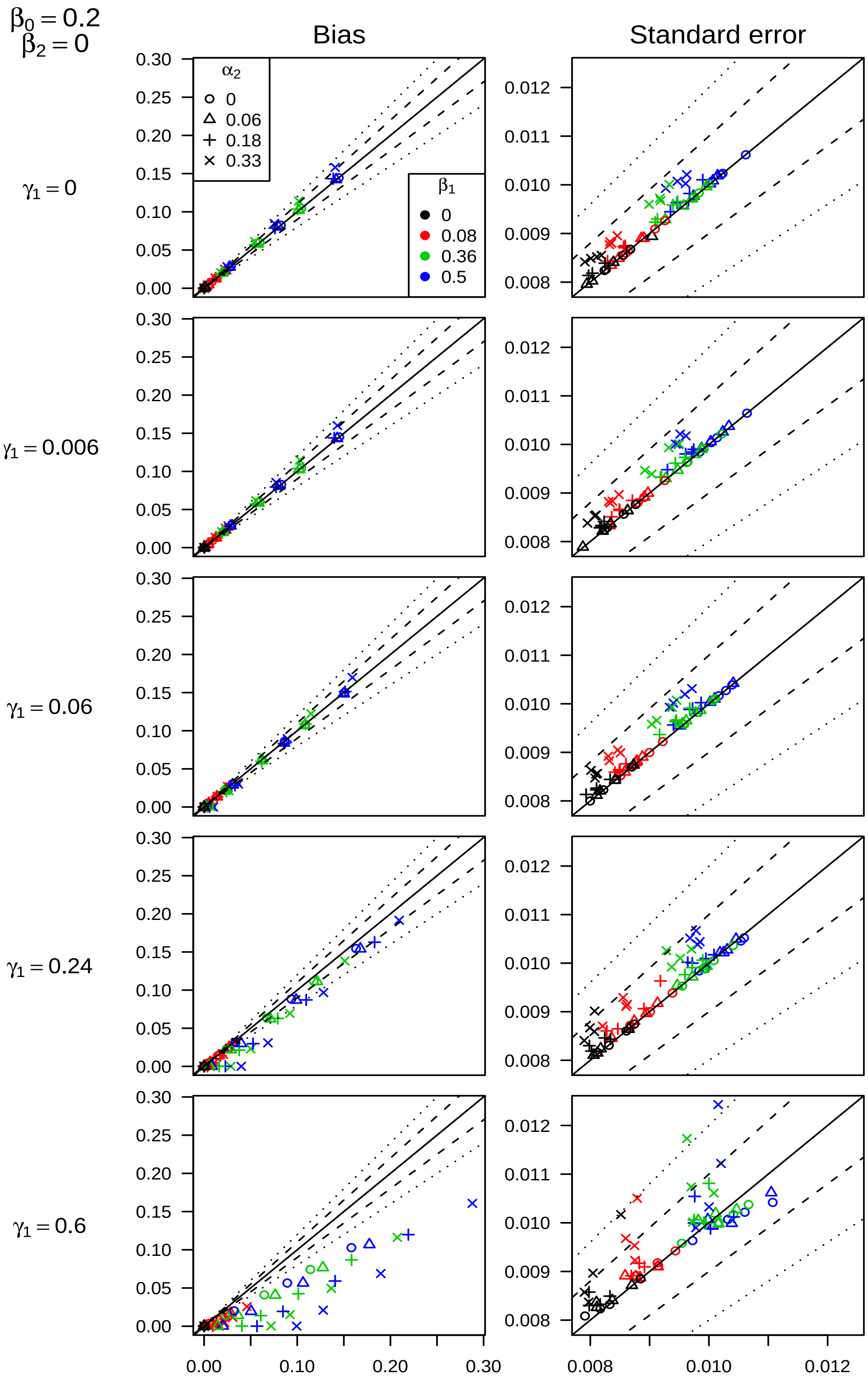
In the following pages, we present the full results of both additive and multiplicative simulation studies. The figures in this section are similar to Figures 4 and 6 in the paper. On the x-axis, we plot the bias (left panel) and standard error (right panel) of $RD_{crude}$. On the y-axis, we plot the bias and standard error of $RD_{cond}$. Each page contains all scenarios for a unique combination of the values of $\beta_0$ and $\beta_2$ and these values are marked in the top left corner of each page. Each row of plots further distinguishes the values of $\gamma_1$, marked to the left of each row. Within each plot, results for all values of $\alpha_1$, $\alpha_2$, and $\beta_1$ are presented, but the values of $\alpha_1$ are not differentiated. The solid diagonal marks equality. Dashed lines represent a 10% increase or 10% decrease, and dotted lines represent a 20% increase or decrease. Web Figures 1-4 are from the additive simulations, and Web Figures 5-10 are from the multiplicative simulations.



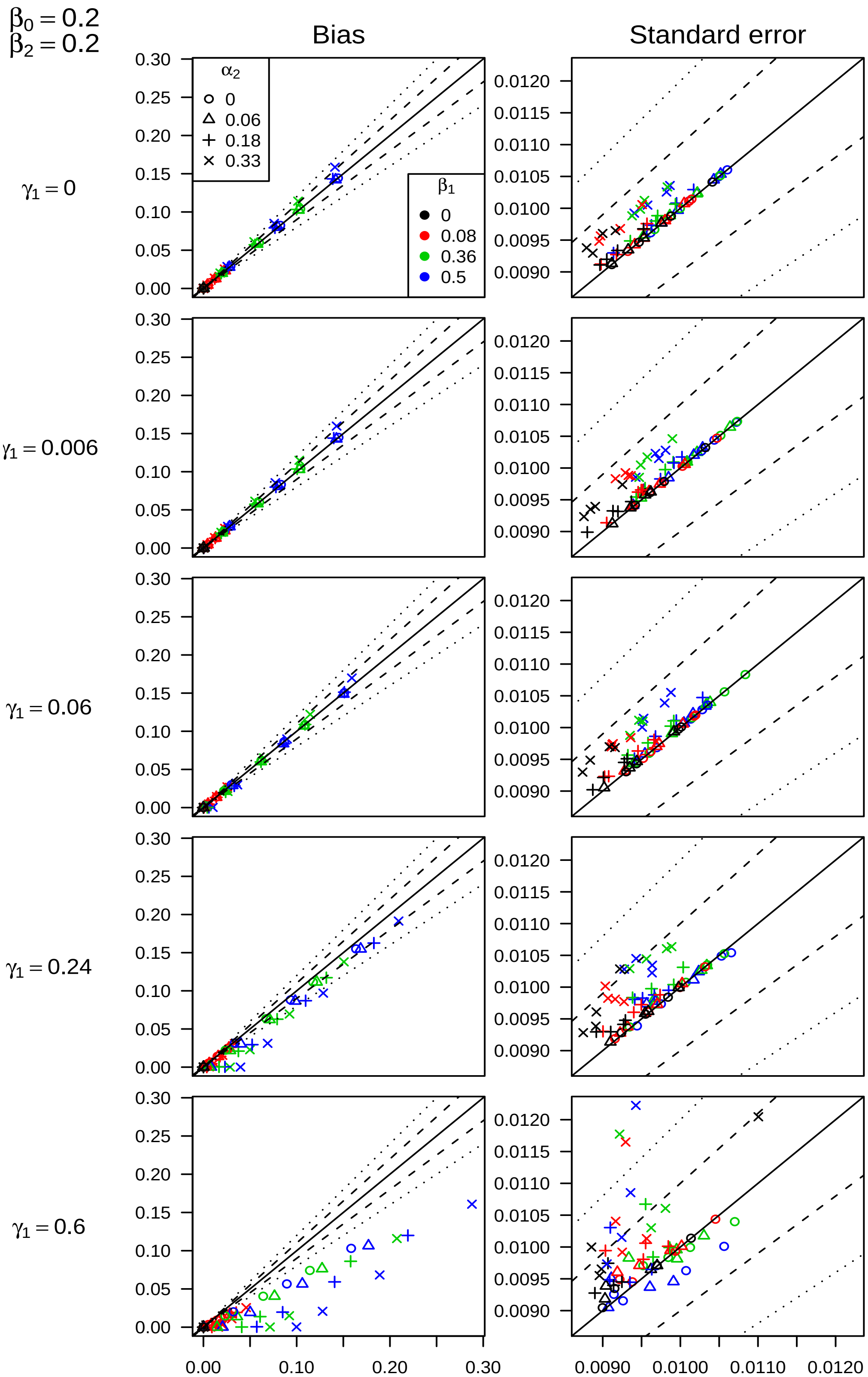

Web Figure 1: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the additive simulations with $\beta_0 = 0.01$ and $\beta_2 = 0$.



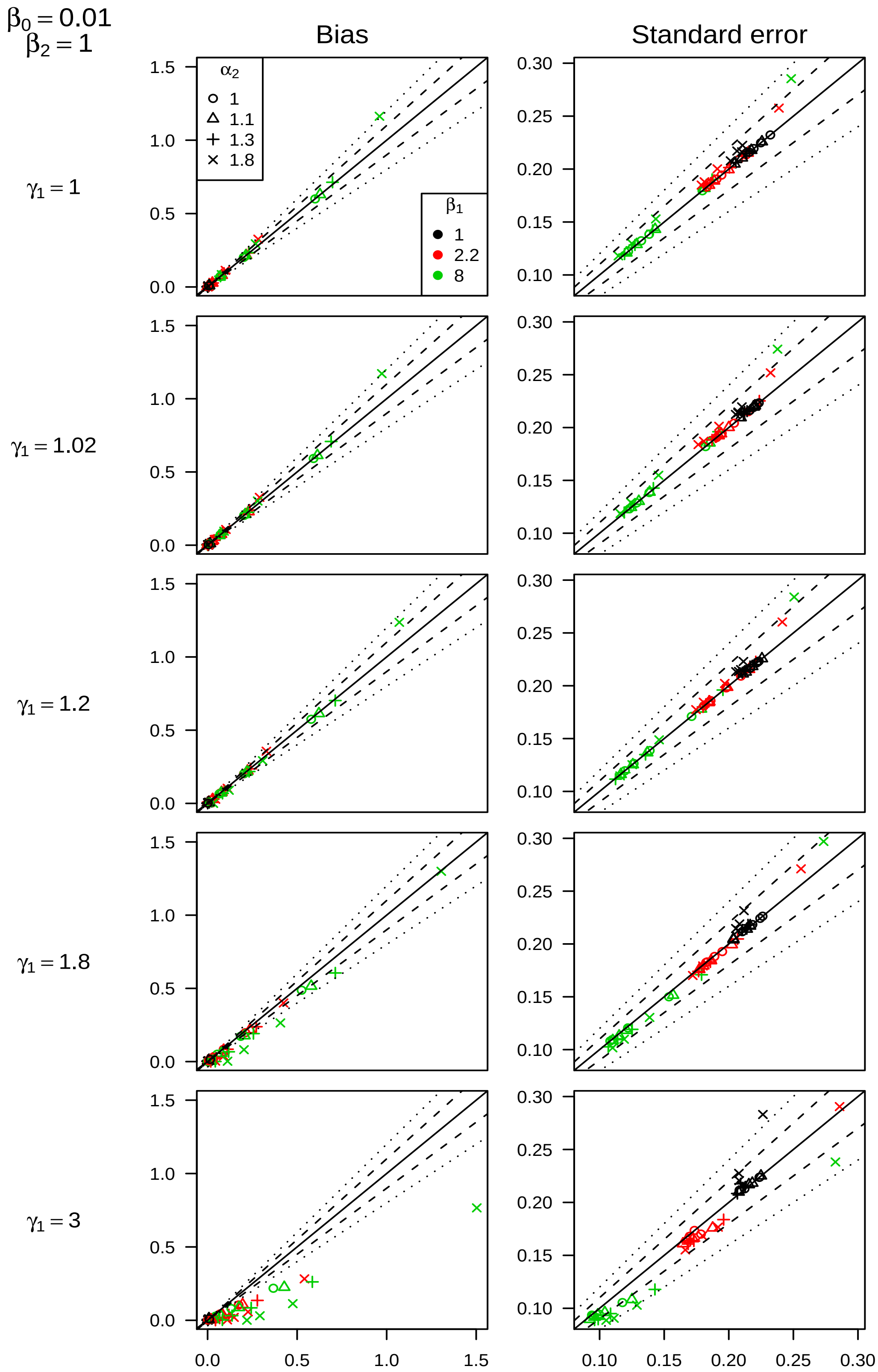

Web Figure 2: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the additive simulations with $\beta_0 = 0.01$ and $\beta_2 = 0.2$.



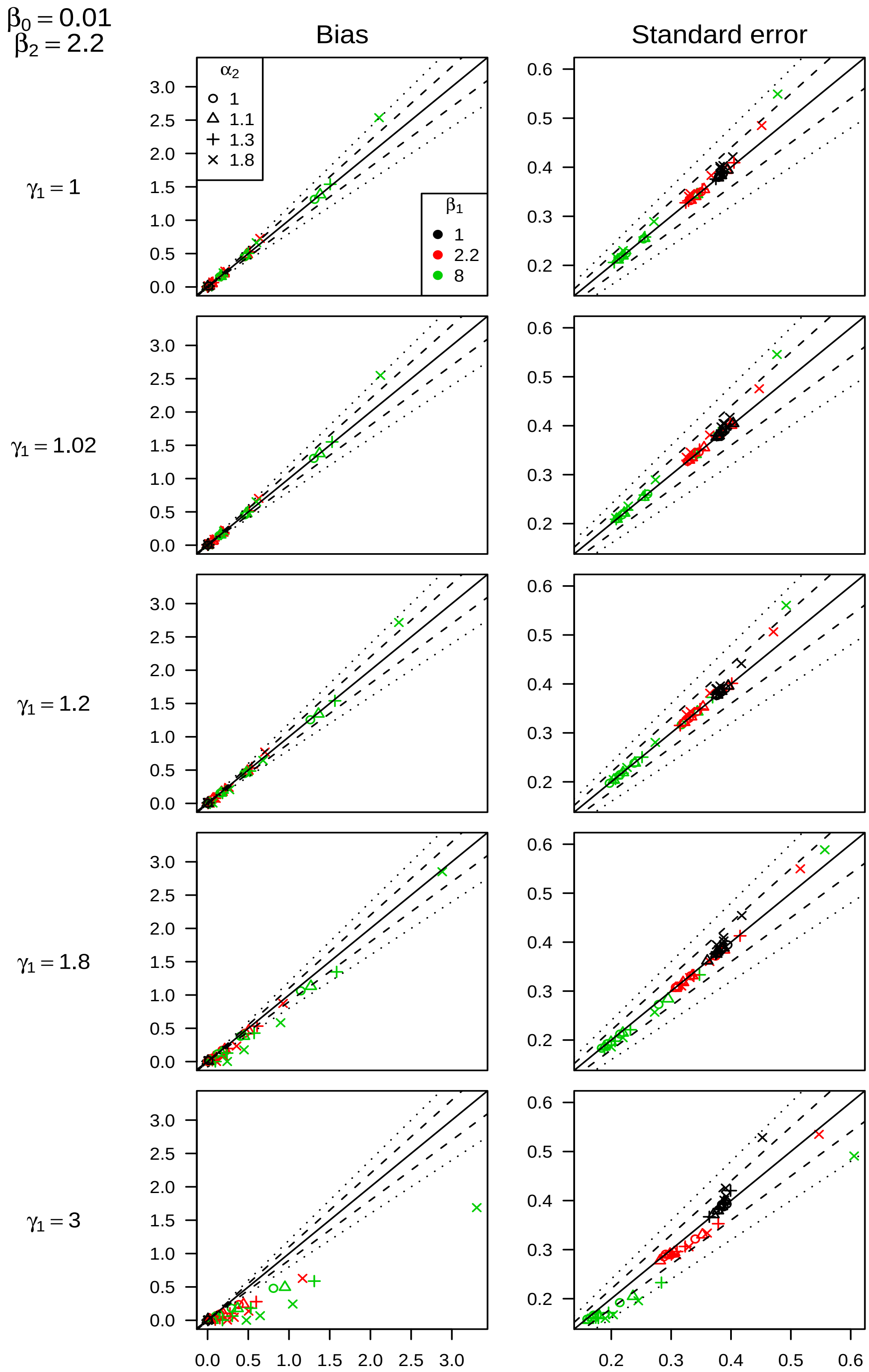

Web Figure 3: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the additive simulations with $\beta_0 = 0.2$ and $\beta_2 = 0$.



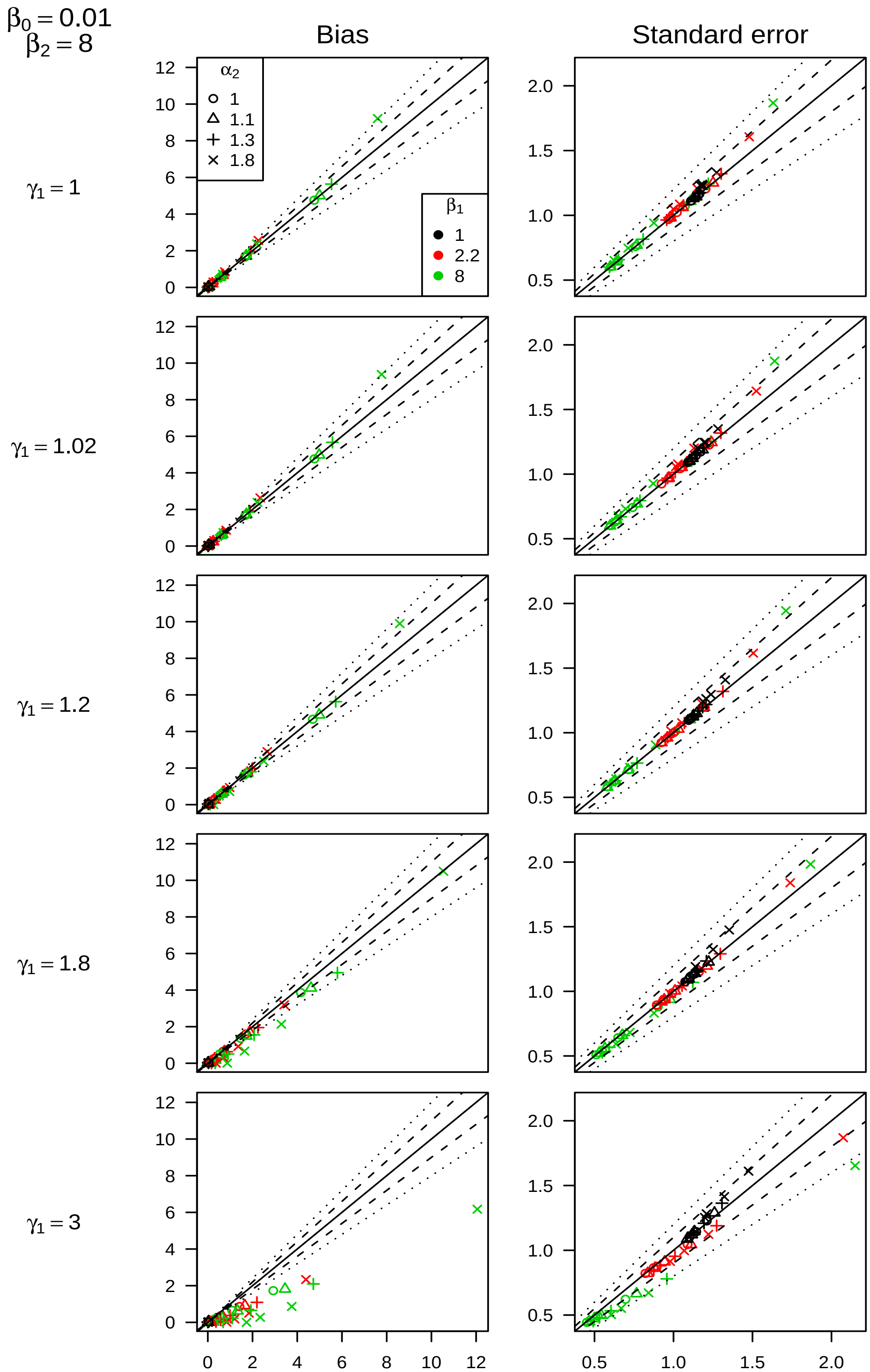

Web Figure 4: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the additive simulations with $\beta_0 = 0.2$ and $\beta_2 = 0.2$.



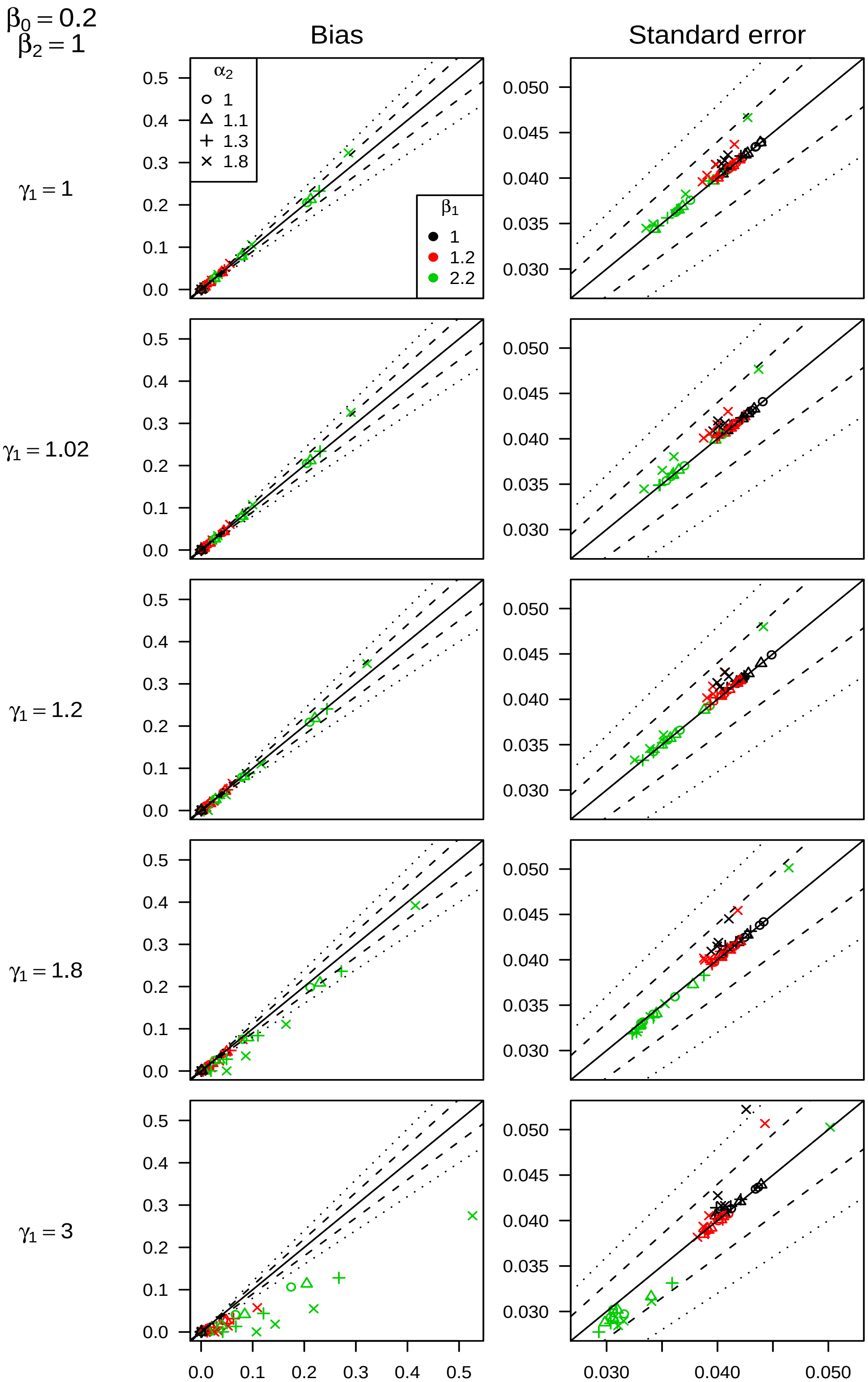

Web Figure 5: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the multiplicative simulations with $\beta_0 = 0.01$ and $\beta_2 = 1$.



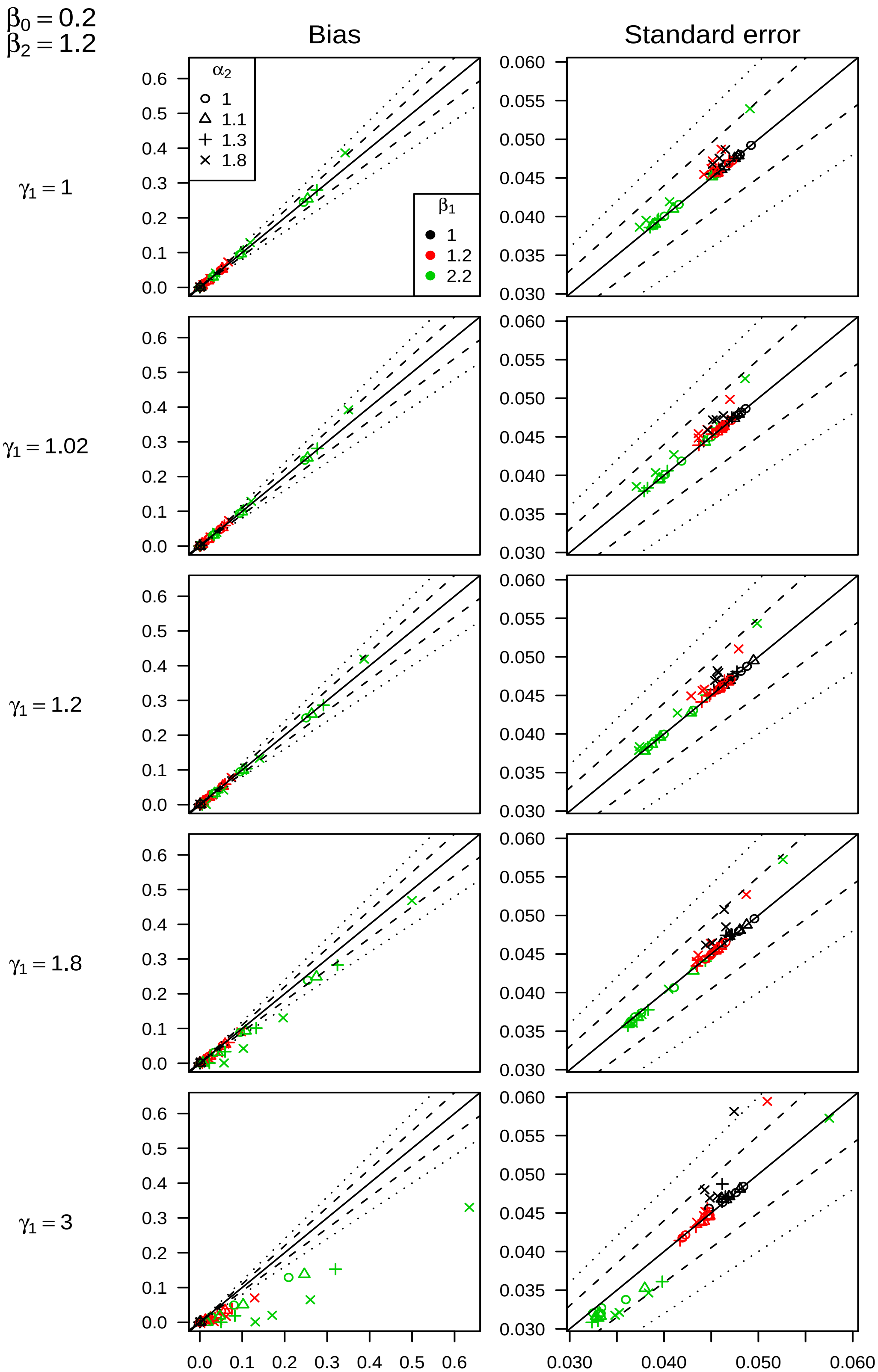

Web Figure 6: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the multiplicative simulations with $\beta_0 = 0.01$ and $\beta_2 = 2.2$.



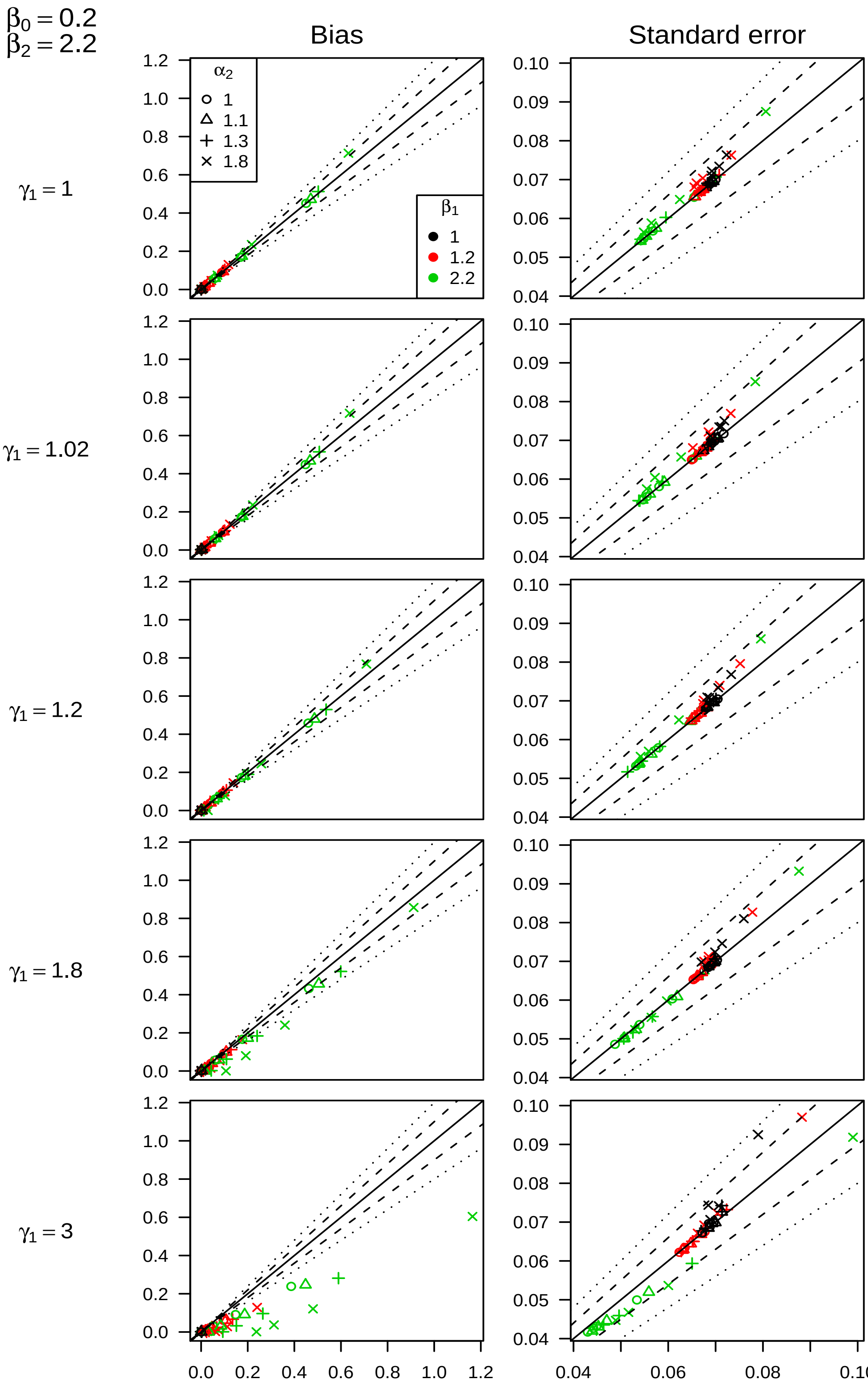

Web Figure 7: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the multiplicative simulations with $\beta_0 = 0.01$ and $\beta_2 = 8$.




Web Figure 8: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the multiplicative simulations with $\beta_0 = 0.2$ and $\beta_2 = 1$.




Web Figure 9: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the multiplicative simulations with $\beta_0 = 0.2$ and $\beta_2 = 1.2$.




Web Figure 10: The bias and standard error of exposure effect estimators with and without conditioning on $Z$. Each point represents one simulation scenario in the multiplicative simulations with $\beta_0 = 0.2$ and $\beta_2 = 2.2$.